

## DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets <sup>7</sup> : G06F 17/30		A1	(11) Numéro de publication internationale: <b>WO 00/60498</b> (43) Date de publication internationale: 12 octobre 2000 (12.10.00)
<p>(21) Numéro de la demande internationale: PCT/FR99/02441 (22) Date de dépôt international: 11 octobre 1999 (11.10.99)</p> <p>(30) Données relatives à la priorité: 99/04130 2 avril 1999 (02.04.99) FR <i>02 Dec 00/20 mby</i></p> <p>(71) Déposant (<i>pour tous les Etats désignés sauf US</i>): BULL S.A. [FR/FR]; 68, route de Versailles, F-78430 Louveciennes (FR).</p> <p>(72) Inventeur; et (75) Inventeur/Déposant (<i>US seulement</i>): NIVELET, Bernard [FR/FR]; 6, allée des Boutons d'Or, F-78310 Maurepas (FR).</p> <p>(74) Mandataire: LEROUX, Jean-Philippe; Bull S.A., 68, route de Versailles (P.C.:58D20), F-78434 Louveciennes Cedex (FR).</p>		<p>(81) Etats désignés: JP, US, brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Publiée <i>Avec rapport de recherche internationale.</i></p>	
<p>(54) Titre: METHOD FOR PRECONDITIONING AND ENCODING A DATA TABLE AND METHOD FOR THE IMPLEMENTATION OF TABLE REQUESTS ON A VECTORAL PROCESSOR</p> <p>(54) Titre: PROCEDE DE PRECONDITIONNEMENT ET DE CODAGE D'UNE TABLE DE DONNEES, ET PROCEDE DE MISE EN OEUVRE DE REQUETES TABULAIRES SUR UN PROCESSEUR VECTORIEL</p> <p>(57) Abstract</p> <p>The invention relates to a search engine (2) that is implemented by a decisional applications server (1) acting on a relational database (6) containing a set of target articles. The search engine (2) is activated by article selection queries according to given criteria and comprises means (8) for preconditioning the database (6), supplying a preconditioned coded table (10) which is periodically updated at the same time as the relational data base itself to a vectoral feature machine (9) for the treatment thereof. Said search engine also contains means (7) for extracting said target articles which are activated by queries on the basis of the result of the treatment of the table (10) installed in the vectoral feature machine (9) from the relational data base (6). The invention can be used in data warehousing systems.</p>			
<p>8...DATA PRECONDITIONING AND DATABASE CREATION MODULE 12...TABLE COHESION MODULE</p>			

(57) Abrégé

L'invention concerne un moteur de recherche (2) mis en oeuvre par un serveur applicatif décisionnel (1) agissant sur une base de données relationnelle (6) qui contient un ensemble d'articles cible. Le moteur (2) est activé par des requêtes de sélection d'articles sur des critères déterminés et comporte des moyens de préconditionnement (8) de la base (6) fourniissant une table codée (10) préconditionnée, remise à jour périodiquement en même temps que la base relationnelle (6) elle-même, à une machine à capacités vectorielles (9) pour son traitement. Il comporte en outre des moyens d'extraction (7) des articles cible, activés par les requêtes, à partir du résultat du traitement de la table (10) installée dans la machine à capacités vectorielles (9), et de la base de données relationnelle (6). Les applications relèvent notamment des systèmes de data warehousing.

**UNIQUEMENT A TITRE D'INFORMATION**

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yugoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		

## PROCEDE DE PRECONDITIONNEMENT ET DE CODAGE D'UNE TABLE DE DONNEES, ET PROCEDE DE MISE EN OEUVRE DE REQUETES TABULAIRES SUR UN PROCESSEUR VECTORIEL

5 La présente invention concerne un procédé de préconditionnement d'une table de données destinée à être utilisé par un moteur de recherche répondant à des requêtes de sélection d'articles sur des critères déterminés.

Elle concerne également un procédé de recherche d'articles, en réponse à une requête déterminée, dans une table de données et un moteur de recherche 10 agissant sur une table de données contenant un ensemble d'articles cible, activé par des requêtes de sélection d'articles sur des critères déterminés.

Le domaine d'application est celui de l'entrepôt de données plus connu sous la terminologie anglo-saxonne de "data warehousing". Il s'adresse plus particulièrement à de grandes bases de données historiques, relativement 15 stables dans le temps à partir desquelles on veut pouvoir extraire des populations définies par critères avec une sollicitation très fréquente et des temps de réponse les plus faibles possibles.

Typiquement ces bases peuvent contenir plusieurs millions d'articles pouvant comporter chacun plusieurs centaines de champs et les temps de réponse à 20 des requêtes standards peuvent être longs alors que l'utilisateur souhaiterait qu'ils soient de l'ordre de la seconde. Elles sont mises à jour périodiquement, au plus chaque nuit.

Les clients potentiels de ce type de bases sont par exemple, la grande distribution, les banques et les assurances.

25 La grande distribution manipule des bases historiques des caisses et des cartes d'achat pour rechercher des populations cible pour le marketing direct.

Les banques et les assurances manipulent également de telles bases historiques relatives à des ordres client, pour des recherches de populations, clientes potentielles de nouveaux produits, etc.

30 On connaît des solutions reposant sur la mise en œuvre du parallélisme pour l'extraction d'articles sur unités de stockage.

Toutes les solutions connues utilisent un mécanisme de gestion de bases de données relationnelles mises à jour et consultées depuis un environnement réseau. Ce mécanisme est connu sous l'acronyme RDBMS (Relational DataBase Management System).

- 5 Dans un premier type de solution, un moteur de recherche SQL (Search Query Language), complètement propriétaire, est construit sur une architecture hautement parallèle à base de nœuds multiprocesseurs pilotant des disques sur lesquels la base de données est répartie. Les requêtes sont fractionnées sur les différents nœuds, puis sur les processeurs.
- 10 Cette solution a comme principal inconvénient d'avoir un rapport prix/performance qui reste élevé. Ainsi, pour atteindre de hautes performances, les configurations doivent être importantes, donc très coûteuses.  
Un deuxième type de solutions utilise des logiciels de bases de données relationnelles standards sur des machines standards généralement 15 multiprocesseurs.

Dans ce deuxième type de solutions, un moteur de recherche SQL standard met en œuvre le parallélisme élevé suivant les mêmes principes que les solutions du premier type mais avec des variantes d'architecture sur les mécanismes de fractionnement des requêtes et sur la gestion des 20 antémémoires.

Les inconvénients de ce deuxième type de solutions sont les mêmes que ceux du premier type, aggravés d'une perte de performance due à la lourdeur du logiciel qui est une conséquence de sa généralité.

L'invention a notamment pour but de pallier ces inconvénients en fournissant 25 un moteur de recherche assez puissant pour effectuer dans des délais très courts, de l'ordre de quelques secondes, sur des tables de données de grande dimension mais stables dans le temps (mises à jour périodiques, au plus chaque nuit), des requêtes de sélection d'articles sur critères.

A cet effet, l'invention a pour premier objet un procédé de préconditionnement 30 d'une ou plusieurs tables de données d'un serveur applicatif décisionnel, destinée à être traitée par un moteur de recherche répondant à des requêtes

de sélection d'articles sur des critères déterminés, émises par le serveur applicatif décisionnel.

Le procédé selon l'invention consiste à :

- analyser les prédictats contenus dans les champs des articles destinés à remplir la base de données relationnelle en fonction de relations autorisée déterminés ;
- créer une nomenclature des prédictats à partir de cette analyse ;
- coder numériquement les prédictats conformément à la nomenclature, en tenant compte de la nature des prédictats et des relations à mettre en œuvre sur les prédictats dans les requêtes.

Il consiste enfin à présenter les prédictats codés, sous la forme d'une table de valeurs numériques.

L'invention a pour deuxième objet un procédé de recherche d'articles, en réponse à une requête déterminée, dans une table de données, consistant à installer une copie de la table de valeurs numériques obtenue par le procédé précédent, sur une machine à capacité vectorielle opérant le traitement des valeurs numériques du tableau en fonctions de la requête servie par le serveur applicatif décisionnel.

Enfin elle a pour troisième objet un système de recherche mis en œuvre par un serveur applicatif décisionnel comportant une base de données relationnelle contenant un ensemble d'articles cible, et un moteur de recherche couplé au serveur applicatif décisionnel, activé par une requête de sélection d'articles sur des critères déterminés émise par le serveur applicatif décisionnel.

Selon l'invention, le système est caractérisé en ce que le moteur comporte des moyens de préconditionnement des données de la base et d'installation d'une table codée, correspondant à la base, sur une machine à capacités vectorielles, ces moyens comportant :

- des moyens de lecture d'un fichier de données correspondant à la base ;
- des moyens pour constituer une nomenclature des valeurs des champs contenus dans le fichier précédent ;

- des moyens de codage des champs conformément à la nomenclature en tenant compte de la nature des champs et des relations à mettre en œuvre sur les prédictats dans la requête ;
- des moyens d'analyse des requêtes émises par le serveur applicatif décisionnel, en tenant compte de relations autorisées, de contraintes sur les prédictats et de la nomenclature ; et
- des moyens de codage de la requête filtrée, en un ensemble de vecteurs contenant les valeurs à trouver dans les champs selon les relations associées, sous forme d'un fichier d'entrée exploitable par la machine à capacités vectorielles.

Le système comporte en outre des moyens pour extraire en clair les données recherchées dans le fichier résultat obtenu en sortie de la machine à capacités vectorielles à partir de moyens de recherche installés dans le serveur applicatif décisionnel.

15 Des synthèses statistiques peuvent être en outre opérées sur les résultats de la recherche.

L'invention a notamment pour avantage de fournir des temps de réponse très courts, inaccessibles par les techniques RDBMS, et un débit de requêtes élevé.

20 Elle a pour autres avantages d'être transparente pour l'application existante et de n'avoir aucun impact au niveau applicatif.

D'autres avantages et caractéristiques de la présente invention apparaîtront à la lecture de la description qui suit faite en référence aux figures annexées qui représentent :

25 – la figure 1, un schéma de principe d'un système de recherche mettant en œuvre un moteur de recherche selon l'invention;

– la figure 2, un schéma de principe d'un module de préconditionnement et d'installation de la base de données, selon l'invention ; et

– la figure 3, un schéma de principe d'un agent SELECT selon l'invention.

30 Dans ces figures, les éléments homologues sont désignés par les mêmes références numériques.

Le principe de l'invention est décrit ci-après en s'appuyant pour son illustration, sur l'utilisation d'une machine vectorielle connue sous l'appellation de supercalculateur.

Une telle machine est caractérisée, d'une part, par des processeurs possédant plusieurs unités arithmétiques, "pipelines" en terminologie anglo-saxonne et, d'autre part, par une bande passante mémoire suffisante pour assurer l'alimentation de tous les processeurs à chaque top d'horloge.

L'invention n'est cependant pas limitée à ce type de machine et s'applique à toute machine à capacités vectorielles, c'est-à-dire des machines dont les performances se rapprochent de celles des supercalculateurs vectoriels.

En effet, les calculateurs scalaires actuels comportent plusieurs opérateurs arithmétiques et les bandes passantes mémoire progressent grâce à l'usage d'une technologie, connue sous la terminologie anglo-saxonne "crossbar" (autocommutateur). On peut donc envisager dans un avenir proche que les performances des calculateurs scalaires se rapprochent de celles des supercalculateurs vectoriels.

Les supercalculateurs vectoriels offrent dès maintenant une réponse aux demandes toujours grandissantes de performances dans les domaines des sciences et de l'industrie en général.

Les machines vectorielles sont aujourd'hui les seules à pouvoir répondre aux contraintes déjà exprimées dans le préambule de la présente description.

L'idée de base de l'invention est de tirer parti de la puissance exceptionnelle des machines à capacités vectorielles pour effectuer des comparaisons sur les vecteurs numériques, images codées des champs de la table de données.

La transformation en nombres des données de la table à traiter, et la formation d'une nomenclature à partir de ces nombres sont réalisées lors de l'installation de la base de données relationnelle.

Le codage des données en nombres a pour autre effet avantageux de compacter les données de la base. Ainsi par rapport aux solutions de type RDBMS qui manipulent le contenu en clair de chaque champ, le procédé selon l'invention n'agit que sur un nombre représentatif de ce champ.

La table ainsi compactée peut être en général contenue en mémoire (pas d'entrée-sortie disque), ou bien peut être chargée en mémoire par colonne ce qui ne représente que des volumes réduits d'entrées/sorties.

Enfin l'invention offre la possibilité d'adapter le codage au type de requêtes qui 5 seront servies. Elle permet en outre la mise en oeuvre d'une optimisation efficace du traitement.

La figure 1 illustre un schéma de principe d'un système de recherche mettant en œuvre un moteur de recherche selon l'invention.

Le système de recherche comporte, à gauche de la figure, un système 10 applicatif décisionnel 1, représentatif du cas général, délimité par une ligne fermée discontinue, et le moteur de recherche 2, à droite de la figure, délimité par une ligne fermée discontinue.

Le système applicatif décisionnel 1 est couplé à un poste utilisateur (ou client) 3.

15 Le serveur applicatif décisionnel 1 comporte un serveur applicatif 4 générant des requêtes prédéfinies, un RDBMS 5 gérant une base de données 6, et un agent SQL 7 chargé d'analyser les requêtes soumises par le serveur applicatif 4 et, éventuellement, d'extraire les articles cible de la base 6 en s'appuyant sur le RDBMS 5.

20 L'utilisateur (le client) émet, via le serveur applicatif 4, des requêtes correspondant à des caractéristiques des articles cible répondant à des critères déterminés et reçoit du même serveur 4 le résultat des requêtes sous la forme soit d'une liste d'articles satisfaisant les critères soit de synthèses statistiques, ou les deux.

25 Le moteur 2 met en œuvre un module 8 de préconditionnement de la table de données et exploite les ressources d'un supercalculateur 9 pour le traitement d'un copie 10 de la table préconditionnée en vue d'en extraire les articles cible. Le module 8 de préconditionnement de la table de données reçoit les données importées par exemple d'une banque de données 11. Ces données sont 30 organisées sous forme d'un tableau et codées numériquement selon un format

directement exploitable par le supercalculateur 9 et exécutable de façon optimale par les requêtes.

Une copie de cette table est accessible par le supercalculateur 9. Elle réside, par exemple, dans l'espace mémoire du supercalculateur 9 et peut être 5 partitionnée si sa taille dépasse celle de la mémoire disponible.

Le supercalculateur 9 reçoit de l'agent SQL 7 la traduction des requêtes soumises par le serveur applicatif 4 sous forme d'un fichier d'entrée.

Le supercalculateur 9 traite ensuite ce fichier d'entrée selon un programme de recherche déterminé tirant le parti maximum de la puissance des pipelines du 10 supercalculateur 9 en travaillant sur les colonnes de la copie 10 de la table.

A l'issue du traitement, il délivre en sortie sous forme d'un fichier, les résultats du traitement effectué correspondant à la liste des numéros de ligne des articles sélectionnés par la recherche et, éventuellement, aux synthèses statistiques demandées sur les articles trouvés.

15 Si les articles en clair sont demandés, l'agent SQL 7 exploite le fichier résultat pour extraire les articles sélectionnés en clair, à partir de la base de données relationnelle 6.

L'agent SQL 7 transmet ensuite les résultats (articles sélectionnés et/ou synthèses statistiques) sous forme de réponse SQL au serveur applicatif 4 20 émetteur de la requête.

Un module 12 de cohérence de tables, accessible à l'agent SQL 7, contient la liste des identifiants des tables présentes et la nomenclature des prédictats pour chacune d'entre elles.

La figure 2 illustre un schéma de principe du module 8 de préconditionnement 25 et d'installation de la table, selon l'invention, délimité par une ligne fermée discontinue.

Il comporte des premiers moyens 13 de lecture des données importées sur un support quelconque, article par article, à l'entrée du module, et provenant par exemple d'une banque de données 11.

Les articles lus sont ensuite, d'une part, complétés de leur numéro et transmis au système de gestion de base de données relationnelle 5 qui crée la base de données en clair 6 dans le serveur applicatif décisionnel 1.

Il comporte, d'autre part, des deuxièmes moyens 14 qui analysent les prédictats 5 sur les articles en fonction de relations autorisées et de contraintes sur les prédictats.

On donne ci-après deux exemples de contraintes sur les prédictats :

Dans un premier exemple, une colonne de la base de données ne comporte que des valeurs numériques. Dans cet exemple, il n'est pas nécessaire de 10 coder numériquement ce qui est déjà numérique.

Dans un deuxième exemple, une colonne de la base de données ne contient que des mots pour lesquels l'ordre alphabétique sera utilisé dans les recherches. Dans cet exemple, l'analyse du prédicat tiendra compte de cette relation pour le codage numérique du prédicat.(préservation de l'ordre)

15 Des troisièmes moyens 15 codent les valeurs des prédictats issus des deuxièmes moyens 14. Ce codage consiste à remplacer les valeurs des champs par leurs index dans la nomenclature des valeurs possibles.

Des quatrièmes moyens 16 créent une nomenclature des prédictats issus des deuxièmes moyens 14 en fonction du codage par les troisièmes moyens 15.

20 Le module de préconditionnement 8 fournit d'autre part l'identifiant de la base codée.

La table codée, la nomenclature des prédictats et l'identifiant de la base se présentent sous la forme de fichiers, respectivement référencés 10, 17 et 18 sur la figure.

25 La figure 3 illustre un schéma de principe d'un agent SELECT 19 selon l'invention. Il se substitue à l'agent SQL 7 du serveur applicatif décisionnel 1 qui l'héberge.

Il comporte des moyens 20 de transformation des requêtes, délimités par une ligne fermée en pointillés, lesquelles requêtes sont soumises par le serveur 30 applicatif 4 en fonction de la nomenclature des prédictats 17, des contraintes sur les prédictats et des relations autorisées.

Les moyens 20 de transformation comportent des moyens 21 d'analyse de la requête SELECT et des moyens 22 de codage des prédictats.

Les moyens 21 d'analyse de la requête traduisent la requête en un ensemble de vecteurs représentatifs des champs à trouver et des relations mises en œuvre en tenant compte des relations autorisées.

Les vecteurs sont ensuite codés par les moyens de codage des prédictats en fonction de la nomenclature des prédictats, des contraintes sur les prédictats et des relations autorisées.

Il y a autant de vecteurs que de valeurs possibles dans les champs de la table.

10 L'analyse permet de construire, par ailleurs, pour chacun de ces vecteurs un vecteur définissant quel type de comparaison effectuer pour chacune des valeurs de champ.

Les vecteurs sont agencés sous forme d'un fichier d'entrée exploitable par le supercalculateur 9.

15 Un programme de recherche intégré au supercalculateur exécute les comparaisons entre les vecteurs et toutes les lignes du tableau.

Ces comparaisons sont effectuées colonne par colonne.

En cas de coïncidence d'une ligne, son numéro est conservé et la réponse fournie par le supercalculateur à l'agent SQL 7 se présente sous la forme d'un 20 fichier résultat comportant la liste des numéros correspondant aux lignes sélectionnées. C'est à partir de ce fichier que les synthèses statistiques demandées sont calculées.

Un module d'extraction 23 construit ensuite, si elle est demandée, la réponse en clair destinée au serveur applicatif 4 émetteur de la requête, en extrayant de 25 la base de données relationnelle 6, les articles correspondant à la liste des numéros de lignes du fichier résultat du supercalculateur 9 en utilisant le numéro d'article ajouté à la base 6.

L'agent SELECT 19 fournit en outre l'identifiant de la table. C'est le module de cohérence de tables 12 qui contrôle l'identité de la table à traiter en cas de 30 pluralité de tables.

Un agent d'administration 24 est en outre couplé à l'agent SELECT 19 et permet de contrôler l'activité du supercalculateur 9 et gérer les anomalies. Il active en outre le chargement du programme de recherche dans le supercalculateur 9.

REVENDICATIONS

1. Procédé de préconditionnement d'une ou plusieurs tables de données d'un serveur applicatif décisionnel (1), destinée à être traitée par un moteur de recherche (2) répondant à des requêtes de sélection d'articles sur des critères déterminés, émises par le serveur applicatif décisionnel (1), caractérisé en ce qu'il consiste à :
  - analyser (14) les prédictats contenus dans les champs des articles destinés à remplir la base de données relationnelle (6) en fonction de relations autorisées déterminées ;
  - créer (16) une nomenclature (17) des prédictats à partir de cette analyse ;
  - coder (15) numériquement les prédictats conformément à la nomenclature (17), en tenant compte de la nature des prédictats et des relations à mettre en œuvre sur les prédictats dans les requêtes ; et
  - en ce qu'il consiste à présenter les prédictats codés, sous la forme d'une table (10) de valeurs numériques.
2. Procédé selon la revendication 1, caractérisé en ce que le codage consiste à remplacer les valeurs des prédictats par leur index dans la nomenclature des valeurs possibles.
3. Procédé selon la revendication 1, caractérisé en ce que le codage compacte les données.
- 25 4. Procédé selon l'une quelconque des revendications 1 à 3, caractérisé en ce que le codage prend en compte le type de requête servie.
5. Procédé de recherche d'articles, en réponse à une requête déterminée, dans une table de données, caractérisé en ce qu'il consiste à installer une copie (10) de la table de valeurs numériques obtenue par le procédé selon 30 l'une quelconque des revendications 1 à 4, sur une machine à capacité

vectorielle (9) opérant le traitement des valeurs numériques du tableau en fonctions de la requête servie par le serveur applicatif décisionnel (1).

6. Procédé selon la revendication 5, caractérisé en ce que la requête est exprimée par un ou plusieurs vecteurs représentatifs des valeurs recherchées dans un champ, et en ce que le traitement consiste à comparer le ou les vecteurs à toutes les lignes du tableau, colonne par colonne, en conservant pour chaque coïncidence le numéro de la ligne.
- 10 7. Procédé selon la revendication 6, caractérisé en ce qu'il consiste, à partir de l'ensemble des numéros des lignes sélectionnées et de la base de données relationnelle (6) comprenant un champ additionnel contenant le numéro des lignes, à extraire de la base de données relationnelle (6) les articles recherchés en clair dont les numéros correspondent, en réponse à une requête.
- 15 8. Procédé selon les revendications 6 ou 7, caractérisé en ce qu'il consiste à exprimer les résultats du traitement sous forme statistique dont une synthèse est fournie en réponse à une requête.
- 20 9. Procédé selon l'une quelconque des revendications 5 à 8, caractérisé en ce que la machine à capacités vectorielles (9) est un supercalculateur.
- 25 10. Système de recherche mis en œuvre par un serveur applicatif décisionnel (1) comportant une base de données relationnelle (6) contenant un ensemble d'articles cible, et un moteur de recherche (2) couplé au serveur applicatif décisionnel (1), activé par une requête de sélection d'articles sur des critères déterminés émise par le serveur applicatif décisionnel (1), caractérisé en ce que le moteur (2) comporte des moyens (8) de préconditionnement des données de la base (6) et d'installation d'une table codée (10), correspondant à

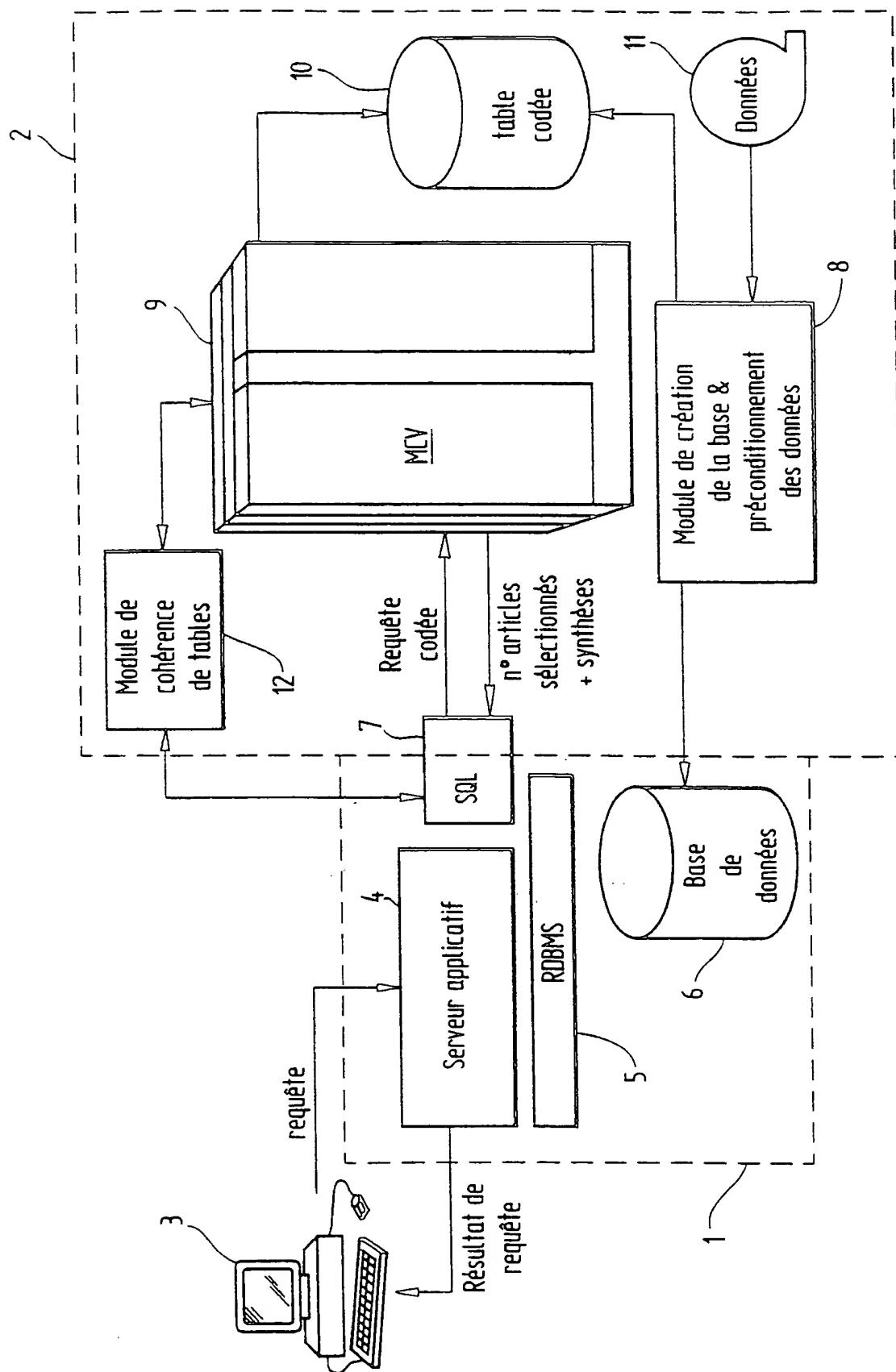
la base (6), sur une machine à capacités vectorielles (9), ces moyens (8) comportant :

- des moyens (13) de lecture d'un fichier de données correspondant à la base ;
- 5 – des moyens (16) pour constituer une nomenclature (17) des valeurs des champs contenus dans le fichier ;
- des moyens (15) de codage des champs conformément à la nomenclature (17) en tenant compte de la nature des champs et des relations à mettre en œuvre sur les prédicts dans la requête ;
- 10 – des moyens (21) d'analyse des requêtes émises par le serveur applicatif décisionnel (1), en tenant compte de relations autorisées, des contraintes sur les prédicts et de la nomenclature (17) ; et
- des moyens (22) de codage de la requête filtrée, en un ensemble de vecteurs contenant les valeurs à trouver dans les champs selon les relations
- 15 associées, sous forme d'un fichier d'entrée exploitable par la machine à capacités vectorielles (9).

11. Système selon la revendication 10, caractérisé en ce qu'il comporte en outre des moyens (23) pour extraire en clair les données recherchées dans le 20 fichier résultat obtenu en sortie de la machine à capacités vectorielles (9) à partir de moyens de recherche installés dans le serveur applicatif décisionnel (1).

12. Système selon l'une quelconque des revendications 10 et 11, caractérisé 25 en ce qu'il comporte en outre, un agent d'administration (24) surveillant l'activité de la machine à capacités vectorielles, gérant les anomalies, et activant les moyens de recherche sur la machine à capacités vectorielles (9).

**This Page Blank (uspto)**

**FIG. 1**

**This Page Blank (uspto)**

2 / 3

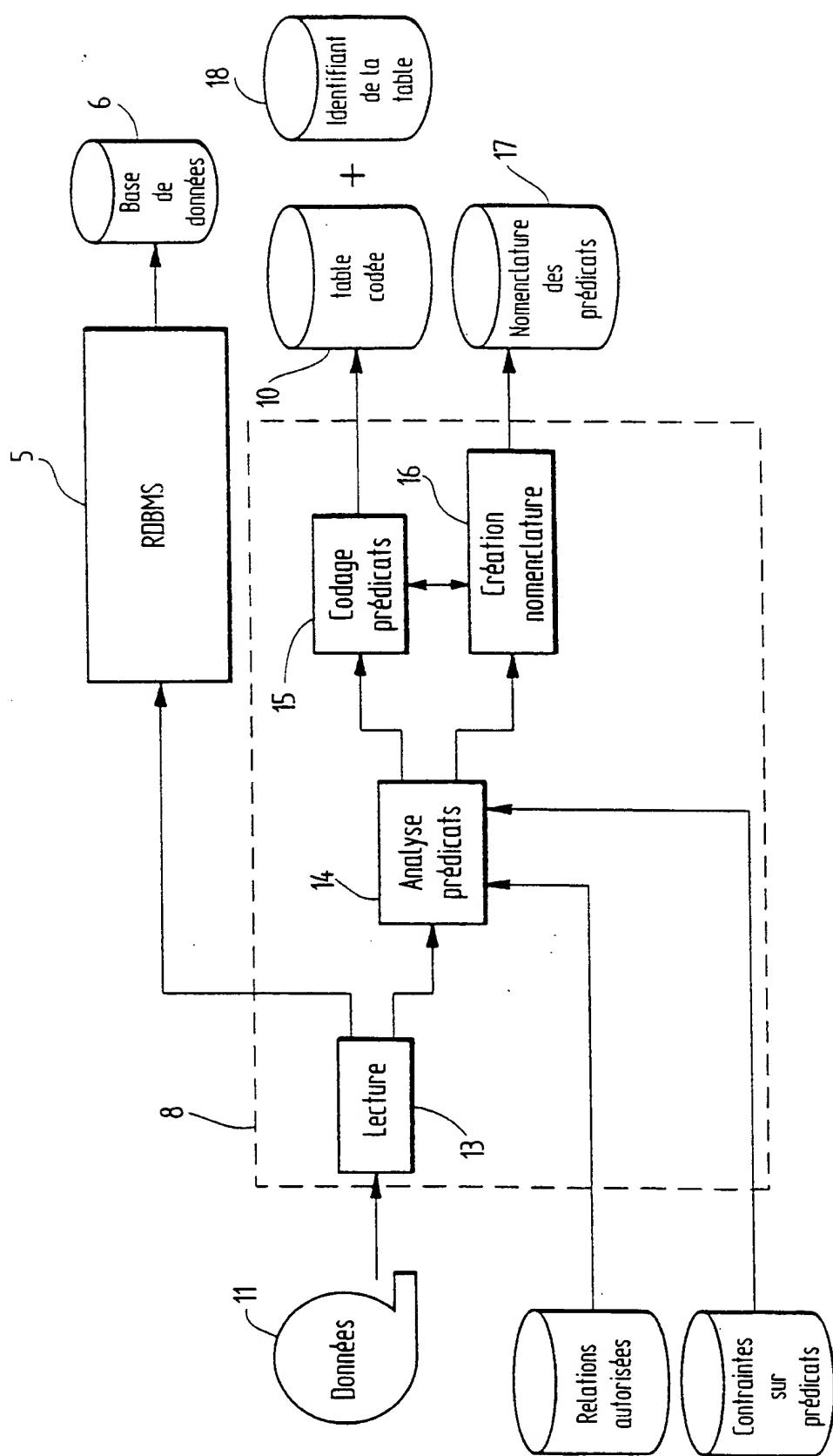


FIG. 2

**This Page Blank (uspto)**

3 / 3

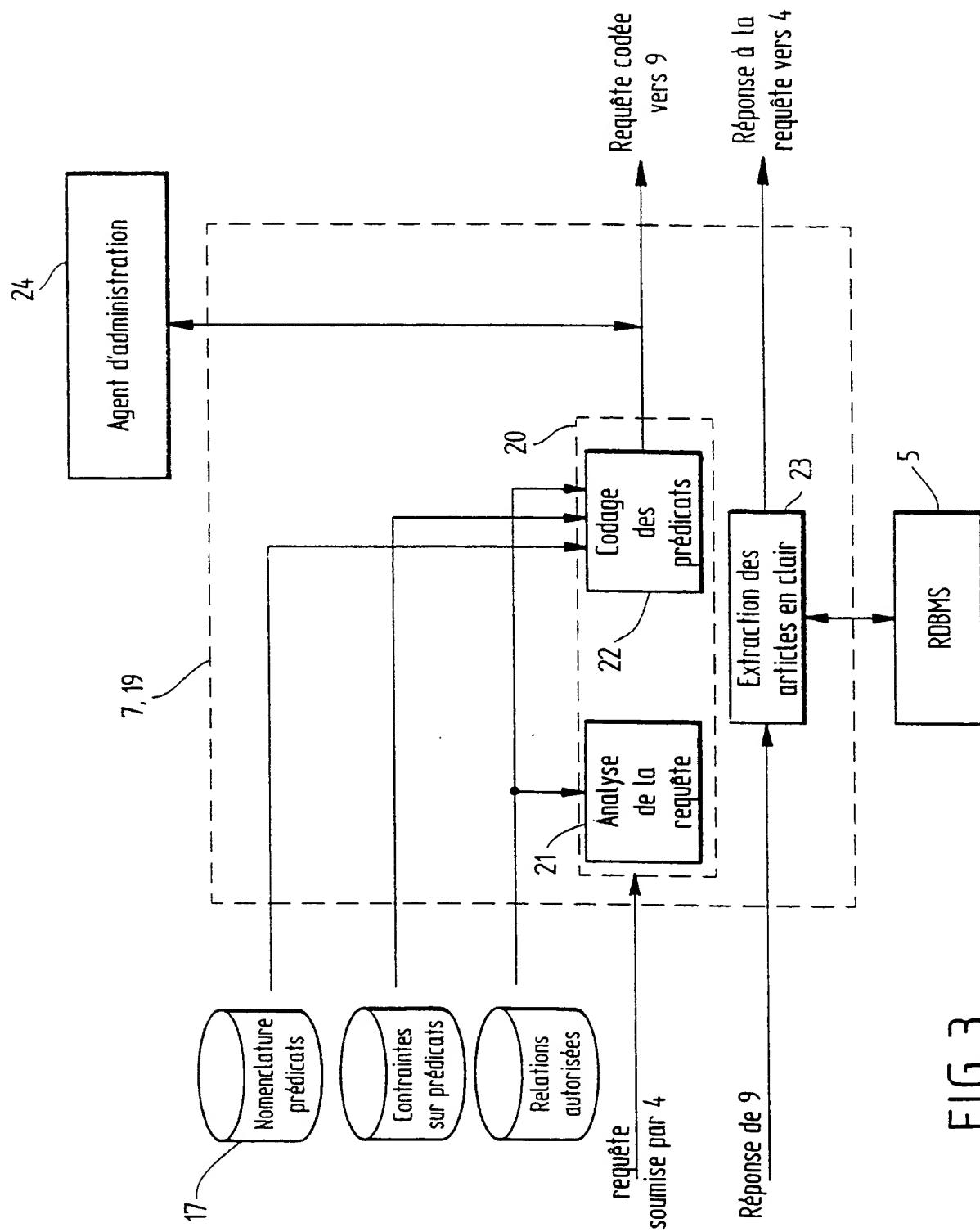


FIG. 3

**This Page Blank (uspto)**

# INTERNATIONAL SEARCH REPORT

Internat'l Application No  
PCT/FR 99/02441

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category <sup>a</sup>	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>SHUN'ICHI TORII ET AL: "ACCELERATING NON-NUMERICAL PROCESSING BY AN EXTENDED VECTOR PROCESSOR" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING, US, WASHINGTON, IEEE COMP. SOC. PRESS, vol. CONF. 4, 1 July 1988 (1988-07-01), page 194-201 XP000124091 page 195, left-hand column, line 6 -page 196, left-hand column, line 8; figures 1-4</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/-</p>	1-7, 9-11

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

<sup>a</sup> Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority, claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

1 December 1999

Date of mailing of the international search report

14/12/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Deane, E

## INTERNATIONAL SEARCH REPORT

Intern: ai Application No  
PCT/FR 99/02441

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOJIMA K ET AL: "A RELATIONAL DATABASE SYSTEM ARCHITECTURE BASED ON A VECTOR PROCESSING METHOD" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING, US, WASHINGTON, IEEE COMP. SOC. PRESS, vol. CONF. 3, 1987, page 182-189 XP000757760 page 185, left-hand column, line 36 -page 186, left-hand column, line 3; figures 5-7 ---	1-7, 9-11
X	ANONYMOUS: "Query Processing With Existing Vector Feature Machines" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 5A, October 1989 (1989-10), pages 305-306, XP002124414 New York, US the whole document ---	1, 10
X	US 4 644 471 A (KOJIMA KEIJI ET AL) 17 February 1987 (1987-02-17) column 1, line 51 - line 68 column 2, line 66 -column 3, line 55; figures 1-4 -----	1, 10

# INTERNATIONAL SEARCH REPORT

## Information on patent family members

Intern. Appl. Application No  
PCT/FR 99/02441

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US 4644471	A 17-02-1987	JP 1998039	C	08-12-1995
		JP 7024036	B	15-03-1995
		JP 60134945	A	18-07-1985
		US 4785400	A	15-11-1988

This Page Blank (uspto)

# RAPPORT DE RECHERCHE INTERNATIONALE

Demar. Internationale No  
PCT/FR 99/02441

A. CLASSEMENT DE L'OBJET DE LA DEMANDE  
CIB 7 G06F17/30

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)  
CIB 7 G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>SHUN'ICHI TORII ET AL: "ACCELERATING NON-NUMERICAL PROCESSING BY AN EXTENDED VECTOR PROCESSOR" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING, US, WASHINGTON, IEEE COMP. SOC. PRESS, vol. CONF. 4, 1 juillet 1988 (1988-07-01), page 194-201 XP000124091 page 195, colonne de gauche, ligne 6 -page 196, colonne de gauche, ligne 8; figures 1-4</p> <p>---</p> <p>-/-</p>	1-7, 9-11

Voir la suite du cadre C pour la fin de la liste des documents

Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent  
"E" document antérieur, mais publié à la date de dépôt international ou après cette date  
"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)  
"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens  
"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

Date d'expédition du présent rapport de recherche internationale

1 décembre 1999

14/12/1999

Nom et adresse postale de l'administration chargée de la recherche internationale  
Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Deane, E

RAPPORT DE RECHERCHE INTERNATIONALE

Demar. Internationale No  
PCT/FR 99/02441

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>KOJIMA K ET AL: "A RELATIONAL DATABASE SYSTEM ARCHITECTURE BASED ON A VECTOR PROCESSING METHOD"          PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING, US, WASHINGTON, IEEE COMP. SOC. PRESS,          vol. CONF. 3, 1987, page 182-189          XP000757760          page 185, colonne de gauche, ligne 36          -page 186, colonne de gauche, ligne 3;          figures 5-7          ---</p>	1-7,9-11
X	<p>ANONYMOUS: "Query Processing With Existing Vector Feature Machines"          IBM TECHNICAL DISCLOSURE BULLETIN,          vol. 32, no. 5A, octobre 1989 (1989-10),          pages 305-306, XP002124414          New York, US          1e document en entier          ---</p>	1,10
X	<p>US 4 644 471 A (KOJIMA KEIJI ET AL)          17 février 1987 (1987-02-17)          colonne 1, ligne 51 - ligne 68          colonne 2, ligne 66 - colonne 3, ligne 55;          figures 1-4          -----</p>	1,10

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demai nternationale No

PCT/FR 99/02441

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)		Date de publication
US 4644471	A	17-02-1987	JP 1998039 C	08-12-1995
			JP 7024036 B	15-03-1995
			JP 60134945 A	18-07-1985
			US 4785400 A	15-11-1988

This Page Blank (uspto)